3526

Improved car detection performance on highways based on YOLOv8

Sutikno, Aris Sugiharto, Retno Kusumaningrum, Helmie Arif Wibawa

Department of Informatics, Faculty of Sciences and Mathematics, University of Diponegoro, Semarang, Indonesia

Article Info

Article history:

Received Dec 15, 2023 Revised Mar 12, 2024 Accepted Mar 28, 2024

Keywords:

Autonomous vehicles Car detection Computer vision Mean average precision YOLOv8 You only look once

ABSTRACT

Car detection on the road through computer vision is crucial for improving safety, as it plays an essential role in spotting nearby vehicles and preventing fatal accidents. Additionally, car detection significantly contributes to the advancement of autonomous vehicles. Previous explorations of car detection using YOLOv5 have revealed weaknesses regarding its resulting mean average precision (mAP). This scenario led to the development of a more advanced version of you only look once (YOLO), namely YOLOv8. Consequently, this study aimed to adopt YOLOv8 for automatic car detection on the road. YOLOv8 is proven to perform better than the previous version. A dataset comprising video frame images was captured on the highway in Semarang, Indonesia. The experiment results indicated that the proposed approach achieved impressive precision, recall, and mAP values, reaching 94.1%, 98.2%, and 98.8%, respectively. The proposed approach enhanced mAP and training time when compared with YOLOv5. Therefore, it was concluded that the proposed method was better suited for real-time car detection.

This is an open access article under the <u>CC BY-SA</u> license.



Corresponding Author:

Sutikno

Department of Informatics, Faculty of Sciences and Mathematics, University of Diponegoro Prof. Jacob Rais Street, Tembalang, Banyumanik, Semarang, Indonesia

Email: sutikno@lecturer.undip.ac.id

1. INTRODUCTION

The annual increase in cars on the road is causing problems such as traffic congestion and safety issues. Car detection technology on the road is crucial to improving safety and transportation efficiency, and its applications have been extended to various study areas. These areas include vehicle identification [1]-[8], car counting [9]-[13], speeding violation detection [14]-[16], and identification of seat belt violations by drivers [17]-[20]. One of the promising solutions for these challenges is computer vision technology. The you only look once (YOLO) algorithm for real-time object detection is a commonly used rapid computational processor.

Several studies have explored car detection using the YOLO algorithm. Sang *et al.* [21] utilized the YOLOv2-vehicle network and BIT-vehicle dataset. In addition, Fei-Fei *et al.* [22] conducted tests on the COCO dataset, comparing various YOLO variants such as YOLOv3, improved YOLOv3, and modified YOLOv3, with modified YOLOv3 showing superior average precision. Wang *et al.* [23] also compared different YOLO variants, including YOLOv2, tiny YOLOv2, tiny YOLOv3, and SPPNet-YOLOv3. The SPPNet-YOLOv3 outperformed the others regarding mean average precision (mAP) [23]. Meanwhile, Jahan *et al.* [24] compared YOLOv3, improved YOLOv3, faster region-based convolutional neural network (R-CNN), and modified YOLOv3. Various YOLO variants were also compared: YOLOv4, YOLOv4 tiny, YOLOv3, and YOLOv3-tiny, revealing that YOLOv4 achieved higher accuracy [25]. Song and Gu [26]

Journal homepage: http://beei.org

introduced YOLOv5 for real-time car detection, comparing it with traditional detection methods, which resulted in fewer false detections. On the other hand, Rafi *et al.* [27] proposed YOLOv5 for detection and tracking, surpassing other models in terms of mAP. However, it is a disadvantage regarding detection process speed. Due to this, a real-time technique that can enhance accuracy and speed detection.

YOLO has been widely used for object detection, including YOLOv3 for waste intensity and person detection in social distance [28], [29], YOLOv4 for fine-grain detection [30], and DenseSPH-YOLOv5 for road damage detection [31]. Moreover, Yung et al. [32] compared some versions of YOLO for detecting safety helmets worn by workers. The findings indicated that YOLOv7 outperformed the others, demonstrating higher mAP. Specifically, YOLOv7 showed better accuracy and speed than other methods. YOLOv8 is even more advanced than its previous version, offering high average accuracy. YOLOv8 has been applied for object detection, including traffic sign detection, which can increase mAP by 14% for YOLOv5 and 13% for YOLOv7 [33]. Therefore, this paper proposed car detection on the highway using YOLOv8. The primary contribution lies in the proposed method's ability to achieve relatively high accuracy to be applied in actual conditions.

2. METHOD

The car detection process outlined in this study passed through several stages, as shown in Figure 1. Typically, it consisted of data annotation, dividing data, training process, and testing process. The dataset comprised video frame images, with the primary goal being detecting car objects in these frames.

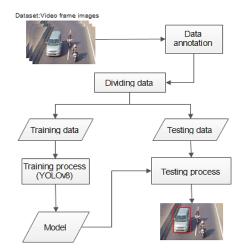


Figure 1. The research method for the car detection

2.1. Dataset

The dataset used consisted of video frame images containing car objects. These frames were extracted from video recordings on the streets of Semarang, Indonesia, particularly Pudak Payung Street, Banyumanik (Figure 2(a)), and Banyumanik Toll Road (Figure 2(b)). Cameras were used to record the video to provide an enhanced output. An example of the frame from the two locations is shown in Figure 2, and the frame had dimensions of 3840×2160 pixels.



Figure 2. Dataset example: (a) road street in Pudak Payung, Banyumanik and (b) toll road in Banyumanik

3528 □ ISSN: 2302-9285

2.2. Data annotation

After acquiring image data, the next step was to annotate each car object in each frame. This annotation was performed by creating bounding boxes around the objects' areas. The total dataset used in this study was 2111 video frame images, split into two sets, namely 80% for training and 20% for testing, consistent with previous explorations [27].

2.3. Training process

We used the YOLOv8 model for the training process. Despite its compact model size, YOLO employed a network to estimate bounding boxes and class, establishing a reputation for achieving accurate object detection [34]. YOLOv8 consists of the backbone and the head, as shown in Figure 3 [35]. The backbone comprised 53 convolution layers with partial connections, which enhanced information flow and feature extraction. However, the head incorporated several convolutions and fully connected layers for object detection.

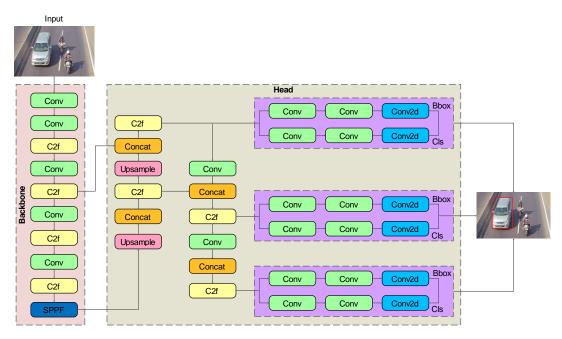


Figure 3. YOLOv8 architecture for car detection

YOLOv8 uses an anchor-free model for classification and regression [36]. This architecture enables each branch to concentrate on its role, improving the model's overall accuracy. The sigmoid function is employed as the activation function. The softmax function is used for class probability, reflecting the likelihood of each item belonging to a specific class.

YOLOv8 uses the complete intersection over union (CIoU) [37] and distribution focal loss (DFL) [38] for bounding-box and classification losses. These losses are utilized to improve the object identification performance. YOLOv8 additionally provides a semantic segmentation model called the YOLOv8-Seg. Instead of the traditional YOLO neck structure, the backbone uses a CSPDarknet53 feature extractor followed by a C2f module that is succeeded by two segmentation heads taught to anticipate semantic segmentation masks for the depicted image.

2.4. Testing process

The entire training process yielded a model used for the testing process. Performance evaluation was based on precision, recall, and mAP, with the added measurement of training time. Precision, recall, and mAP were calculated using (1) to (3) [39], where TP represented true positive detections, FP denoted false positive detections, and FN accounted for ground-truth objects that were not detected.

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$R = \frac{TP}{TP + FN} \tag{2}$$

$$mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \tag{3}$$

APi denotes the average precision for class *i*, and *N* represents the total of classes. The calculation of *AP* is determined by (4). Pi(Rn+1) is computed using (5). Here, $P(\tilde{R})$ represents precision measured at recall \tilde{R} .

$$AP = \sum_{n} (R_{n+1} - R_n) P_i(R_{n+1}) \tag{4}$$

$$P_i(R_{n+1}) = \max_{\tilde{R}: \tilde{R} \ge R_{n+1}} P(\tilde{R}) \tag{5}$$

3. RESULTS AND DISCUSSION

Tesla V100-SXM2-16GB was used for the testing processes, which were conducted using the YOLOv8x model, with 100 epochs and an early stopping set at 15. Testing included various batch sizes, namely 2, 4, 8, and 16, and the results were shown in Table 1. The highest precision achieved was 94.6%, with a batch size 16. The best recall and mAP reached 98.2% and 98.8%, respectively, using a batch size of 8. The shortest time for training recorded was 0.95 hours, with a batch size of 16. Figure 4 shows car detection results in various frame images for visual reference. From this example, it was evident that car objects partially obscured at the frame boundaries went undetected. However, apparent objects in the frame were successfully detected. This limitation could be addressed by including partially obscured car objects in the training data. Additionally, Figures 5(a) to (d) shows the proposed model's precision-recall curve, precision-confidence curve, recall-confidence curve, and F-1-confidence curve. Figure 6 explicitly shows the training and testing performance for the proposed model.

Table 1. Test results with varying batch sizes for car detection

Batch size	Precision (%)	Recall (%)	mAP (%)	Training time (hours)
2	92.5	97.9	97.1	3.24
4	94.4	97.8	98.5	1.22
8	94.1	98.2	98.8	1.69
16	94.6	98.0	97.9	0.95



Figure 4. Example of car detection

3530 □ ISSN: 2302-9285

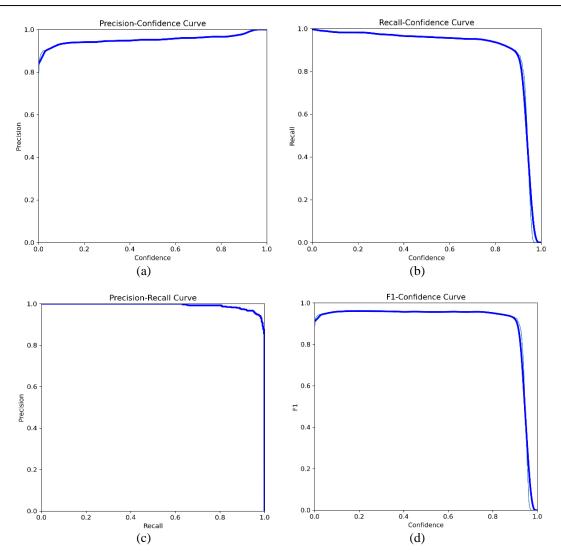


Figure 5. Performance of proposed method: (a) precision-confidence curve, (b) recall-confidence curve, (c) precision-recall curve, and (d) F1-confidence curve

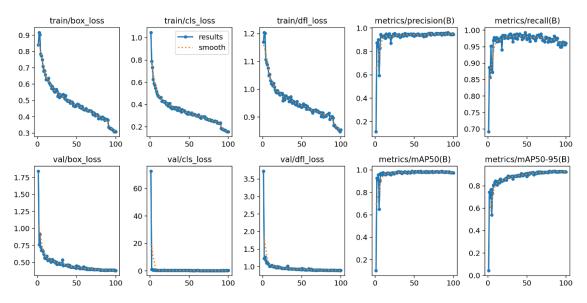


Figure 6. Performance of training and testing in the proposed method

Finally, we compare our method with the previous study, as seen in Table 2. In this comparison, [27] used the YOLOv5l method trained with 100 epochs and an image size of 640×640 pixels. The proposed method outperformed in terms of recall, mAP, and training time. The result showed a 2.2% increase in recall and a 0.4% increment in mAP. In the context of training time, the proposed method was 0.5 times faster than YOLOv5. Therefore, it was concluded that the proposed method offered superior performance and was better suited for real-time applications.

Table 2. Comparison between the proposed model and previous research

Method	Precision (%)	Recall (%)	mAP (%)	Training time (hours)
[27]	96.0	97.0	98.4	3.378
Proposed	94.1	98.2	98.8	1.690

CONCLUSION

This study proposed using YOLOv8 for car detection on the highway. Testing was done with various batch sizes, including 2, 4, 8, and 16. The results showed that the proposed method's precision, recall, and mAP values reached 94.1%, 98.2%, and 98.8%, respectively. Compared to YOLOv5, the proposed method increased recall and mAP values by 2.2% and 0.4%, respectively. Additionally, it reduced training time by half compared to YOLOv5, implying that the proposed approach was better suited for real-time applications. Even though the precession of the proposed method is relatively high, namely 94.1%, the proposed method produces a lower precision vaabovelue than YOLOv5. This precision can be increased by adding a dataset with various types of cars on the road. This study could be extended by implementing license plate detection to identify the license plates on detected cars.

ACKNOWLEDGEMENTS

The authors thank the Ministry of Research and Higher Education of Indonesia for supporting this study via Diponegoro University with contract number 017/E5/PG.02.00.PL/2023. Their support was instrumental in the successful completion of this research.

REFERENCES

- S. Naseer, S. M. A. Shah, S. Aziz, M. U. Khan, and K. Iqtidar, "Vehicle make and model recognition using deep transfer learning and support vector machines," in Proceedings-2020 23rd IEEE International Multi-Topic Conference, INMIC 2020, pp. 4-9, doi: 10.1109/INMIC50486.2020.9318063.
- M. W. Nafi'l, E. M. Yuniarno, and A. Affandi, "Vehicle brands and types detection using mask R-CNN," in Proceedings 2019 International Seminar on Intelligent Technology and Its Application, 2019, pp. 422-427, doi: 10.1109/ISITIA.2019.8937278.
- I. V. Pustokhina et al., "Automatic vehicle license plate recognition using optimal k-means with convolutional neural network for intelligent transportation systems," *IEEE Access*, vol. 8, pp. 92907–92917, 2020, doi: 10.1109/ACCESS.2020.2993008.
- L. Ju-xia, "License plate recognition model research based on the multi-feature technology," TELKOMNIKA Indonesian Journal of Electrical Engineering, vol. 12, no. 4, pp. 2724–2734, 2014, doi: 10.11591/telkomnika.v12i4.4307.
- A. G. S Fakhar, H. M Saad, K. A Fauzan, H. R Affendi, and A. M Aidil, "Development of portable automatic number plate recognition (ANPR) system on Raspberry Pi," International Journal of Electrical and Computer Engineering, vol. 9, no. 3, pp. 1805–1813, 2019, doi: 10.11591/ijece.v9i3.pp1805-1813.
- N. L. Yaacob, A. A. Alkahtani, F. M. Noman, A. W. M. Zuhdi, and D. Habeeb, "License plate recognition for campus auto-gate system," Indonesian Journal of Electrical Engineering and Computer Science, vol. 21, no. 1, pp. 128-136, 2021, doi: 10.11591/ijeecs.v21.i1.pp128-136.
- C. Tu and S. Du, "A hierarchical RCNN for vehicle and vehicle license plate detection and recognition," International Journal of
- Electrical and Computer Engineering, vol. 12, no. 1, pp. 731–737, 2022, doi: 10.11591/ijece.v12i1.pp731-737.

 D. Islam, T. Mahmud, and T. Chowdhury, "An efficient automated vehicle license plate recognition system under image processing," Indonesian Journal of Electrical Engineering and Computer Science, vol. 29, no. 2, pp. 1055-1062, 2023, doi: 10.11591/ijeecs.v29.i2.pp1055-1062.
- M. A. Abdelwahab, "Accurate vehicle counting approach based on deep neural networks," in Proceedings of 2019 International Conference on Innovative Trends in Computer Engineering, 2019, pp. 1–5, doi: 10.1109/ITCE.2019.8646549.
- [10] Q. Chen, N. Huang, J. Zhou, and Z. Tan, "An SSD Algorithm based on vehicle counting method," in Chinese Control Conference, CCC, 2018, pp. 7673–7677, doi: 10.23919/ChiCC.2018.8483037.
- [11] C. M. Tsai, F. Y. Shih, and J. W. Hsieh, "Real-time vehicle counting by deep-learning networks," in Proceedings- International Conference on Machine Learning and Cybernetics, 2022, pp. 175-181, doi: 10.1109/ICMLC56445.2022.9941299.
- [12] J. M. Anil, L. Mathews, R. Renji, R. M. Jose, and S. Thomas, "Vehicle counting based on convolution neural network," in Proceedings of the 7th International Conference on Intelligent Computing and Control Systems, 2023, pp. 695–699, doi: 10.1109/ICICCS56967.2023.10142302.
- [13] M. A. Marzouk and A. A. El Azeem, "Vehicles detection and counting based on internet of things technology and video processing techniques," IAES International Journal of Artificial Intelligence, vol. 11, no. 2, pp. 405-413, 2022, doi: 10.11591/ijai.v11.i2.pp405-413.
- [14] J. Timofejevs, A. Potapovs, and M. Gorobetz, "Algorithms for computer vision based vehicle speed estimation sensor," in 63rd Annual International Scientific Conference on Power and Electrical Engineering of Riga Technical University, 2022, pp. 1-6,

- doi: 10.1109/RTUCON56726.2022.9978802.
- [15] W. Jianping, L. Zhaobin, L. Jinxiang, G. Caidong, S. Maoxin, and T. Fangyong, "An algorithm for automatic vehicle speed detection using video camera," in *Proceedings of 2009 4th International Conference on Computer Science and Education*, 2009, pp. 193–196, doi: 10.1109/ICCSE.2009.5228496.
- [16] A. Lad, P. Kanaujia, Soumya, and Y. Solanki, "Computer vision enabled adaptive speed limit control for vehicle safety," in Proceedings - 2021 1st IEEE International Conference on Artificial Intelligence and Machine Vision, 2021, pp. 4–8, doi: 10.1109/AIMV53313.2021.9670944.
- [17] A. Upadhyay, B. Sutrave, and A. Singh, "Real time seatbelt detection using YOLO deep learning model," in 2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), 2023, pp. 1–6, doi: 10.1109/SCEECS57921.2023.10063114.
- [18] J. Albesa and M. Gasulla, "Seat occupancy and belt detection in removable seats via inductive coupling," in IEEE 74th Vehicular Technology Conference, 2011, pp. 1–5, doi: 10.1109/VETECF.2011.6093145.
- [19] B. Zhou, L. Chen, J. Tian, and Z. Peng, "Learning-based seat belt detection in image using salient gradient," in *Proceedings of the 2017 12th IEEE Conference on Industrial Electronics and Applications*, 2017, pp. 547–550, doi: 10.1109/ICIEA.2017.8282904.
- [20] Z. Wang and Y. Ma, "Detection and recognition of stationary vehicles and seat belts in intelligent internet of things traffic management system," *Neural Computing and Applications*, vol. 34, no. 5, pp. 3513–3522, 2022, doi: 10.1007/s00521-021-05870-6.
- [21] J. Sang et al., "An improved YOLOv2 for vehicle detection," Sensors (Switzerland), vol. 18, no. 12, 2018, doi: 10.3390/s18124272.
- [22] L. Fei-Fei, R. Fergus, and P. Perona, "One-shot learning of object categories," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 4, pp. 594–611, 2006, doi: 10.1109/TPAMI.2006.79.
- [23] X. Wang, S. Wang, J. Cao, and Y. Wang, "Data-driven based tiny-YOLOv3 method for front vehicle detection inducing SPP-Net," *IEEE Access*, vol. 8, pp. 110227–110236, 2020, doi: 10.1109/ACCESS.2020.3001279.
- [24] N. Jahan, S. Islam and M. F. A. Foysal, "Real-time vehicle classification using CNN," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1–6, doi: 10.1109/ICCCNT49239.2020.9225623.
- [25] M. A. B. Zuraimi and F. H. K. Zaman, "Vehicle detection and tracking using YOLO and DeepSORT," in *ISCAIE 2021 IEEE 11th Symposium on Computer Applications and Industrial Electronics*, 2021, pp. 23–29, doi: 10.1109/ISCAIE51753.2021.9431784.
- [26] X. Song and W. Gu, "Multi-objective real-time vehicle detection method based on YOLOv5," in 2021 International Symposium on Artificial Intelligence and its Application on Media (ISAIAM), 2021, pp. 142–145, doi: 10.1109/ISAIAM53259.2021.00037.
- [27] M. M. Rafi et al., "Performance analysis of deep learning YOLO models for south asian regional vehicle recognition," International Journal of Advanced Computer Science and Applications, vol. 13, no. 9, pp. 864–873, 2022, doi: 10.14569/IJACSA.2022.01309100.
- [28] F. F. Putra and Y. D. Prabowo, "Low resource deep learning to detect waste intensity in the river flow," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 5, pp. 2724–2732, 2021, doi: 10.11591/eei.v10i5.3062.
- [29] V. S. Sadanand, K. Anand, P. Suresh, P. K. A. Kumar, and P. Mahabaleshwar, "Social distance and face mask detector system exploiting transfer learning," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 6, pp. 6149–6158, 2022, doi: 10.11591/ijece.v12i6.pp6149-6158.
- [30] A. M. Roy, R. Bose, and J. Bhaduri, "A fast accurate fine-grain object detection model based on YOLOv4 deep neural network," Neural Computing and Applications, vol. 34, no. 5, pp. 3895–3921, 2022, doi: 10.1007/s00521-021-06651-x.
- [31] A. M. Roy and J. Bhaduri, "DenseSPH-YOLOv5: An automated damage detection model based on DenseNet and Swin-Transformer prediction head-enabled YOLOv5 with attention mechanism," Advanced Engineering Informatics, vol. 56, p. 102007, 2023, doi: 10.1016/j.aei.2023.102007.
- [32] N. D. T. Yung, W. K. Wong, F. H. Juwono, and Z. A. Sim, "Safety helmet detection using deep learning: implementation and comparative study using YOLOv5, YOLOv6, and YOLOv7," in 2022 International Conference on Green Energy, Computing and Sustainable Technology, GECOST 2022, 2022, pp. 164–170, doi: 10.1109/GECOST55694.2022.10010490.
- [33] F. N. Ortatas and M. Kaya, "Performance Evaluation of YOLOv5, YOLOv7, and YOLOv8 models in traffic sign detection," in *UBMK 2023-Proceedings: 8th International Conference on Computer Science and Engineering*, 2023, pp. 151–156, doi: 10.1109/UBMK59864.2023.10286611.
- [34] J. Solawetz and Francesco, "What is YOLOv8? The ultimate guide," 2023. [Online]. Available: https://blog.roboflow.com/whats-new-in-yolov8/#the-yolov8-annotation-format. (Accessed: January 10, 2024).
- [35] H. T. Ngoc, K. H. Nguyen, H. K. Hua, H. V. N. Nguyen, and L. D. Quach, "Optimizing YOLO performance for traffic light detection and end-to-end steering control for autonomous vehicles in gazebo-ROS2," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 7, pp. 475–484, 2023, doi: 10.14569/IJACSA.2023.0140752.
- [36] J. Terven, D. M. Córdova-Esparza, and J. A. Romero-González, "A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680– 1716, 2023. doi: 10.3390/make5040083.
- [37] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren, "Distance-IoU loss: faster and better learning for bounding box regression," AAAI 2020 - 34th AAAI Conference on Artificial Intelligence, no. 2, pp. 12993–13000, 2020, doi: 10.1609/aaai.v34i07.6999.
- [38] X. Li et al., "Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection," in Advances in Neural Information Processing Systems, vol. 23, pp. 21002–21012, 2020.
- [39] R. Padilla, S. L. Netto, and E. A. B. D. Silva, "A survey on performance metrics for object-detection algorithms," in *International Conference on Systems, Signals, and Image Processing*, 2020, pp. 237–242, doi: 10.1109/IWSSIP48289.2020.9145130.

BIOGRAPHIES OF AUTHORS



Dr. Sutikno © See received a Doctor of Philosophy (Ph.D.) degree in Computer Science from the Faculty of Mathematics and Natural Sciences at the University of Gadjah Mada, Indonesia. Now, he is an Assistant Professor at the Department of Informatics, Faculty of Sciences and Mathematics, University of Diponegoro. His research interests include machine learning, computer vision, and artificial intelligence. He can be contacted at email: sutikno@lecturer.undip.ac.id.



Dr. Aris Sugiharto Preceived a Doctor of Philosophy (Ph.D.) degree in computer science from the Faculty of Mathematics and Natural Sciences at the University of Gadjah Mada, Indonesia. Now, he is an Assistant Professor at the Department of Informatics, Faculty of Sciences and Mathematics, University of Diponegoro. His current research interests include pattern recognition and computer vision. He can be contacted at email: arissugiharto@lecturer.undip.ac.id.



Dr. Retno Kusumaningrum (b) [S] [S] (M'16) received her B.S. degree in mathematics from Universitas Diponegoro, Semarang, Indonesia, in 2003, and her M.CS, and Ph.D. degrees from Universitas Indonesia, Depok, Indonesia in 2010 and 2014, respectively. She is currently a Lecturer at the Department of Informatics, Faculty of Science and Mathematics, University of Diponegoro. Furthermore, she is currently participating in the Laboratory of Intelligent Systems. Her research interests include machine learning, natural language processing, computer vision, pattern recognition, and topic modeling. She is a member of the IEEE Computational Intelligence Society, IEEE Computer Society, and ACM. Dr. Kusumaningrum's awards and honors include the Sandwich-Like scholarship award from the Directorate General of Higher Education of Indonesia for visiting the School of System Engineering, University of Reading, Reading, U.K. as a student visitor in 2012, the Best Paper of the Second International Conference on Informatics and Computational Sciences in 2018, first place for Outstanding Lecturer - Universitas Diponegoro for the Science and Technology Category in 2019, and second place for the Best Paper Award of the Third International Symposium on Advanced Intelligent Informatics in 2020. She can be contacted at email: retno@live.undip.ac.id.

